

What is claimed is:

1. A burst-switching network comprising:

a plurality of source nodes;

a plurality of upstream links coupled to said plurality of source nodes;

5 a plurality of sink nodes;

a plurality of downstream links coupled to said plurality of sink nodes;

a plurality of core nodes, at least one of said plurality of core nodes is coupled to a subset of said plurality of upstream links and a subset of said plurality of downstream links and has a plurality of space switches, each space switch having a slave controller; and

10

a plurality of master controllers in each core node, one said master controller associated with each of said plurality of space switches in each of said plurality of core nodes and a designated one of said master controllers in a core node functions as a core-node controller, said core-node controller communicatively connecting to each of said master controllers,

15

said core-node controller operable to:

receive control data from at least one of said plurality of source nodes;

divide said control data among said master controllers; and

instruct each master controller to generate a burst-switching schedule for a space switch associated with said each master controller, communicate said schedule to a respective edge node, and transmit instructions based on said schedule to a slave controller of said space switch after a pre-calculated delay period.

20

2. The network of claim 1 wherein each of said plurality of source nodes is paired with a corresponding one of said plurality of sink nodes to form a plurality of network edge nodes.

- 25 3. The network of claim 2 wherein at least one of said plurality of space switches is an optical switch.

4. The network of claim 3 wherein any of said master controllers can be designated to function as a core-node controller.

5. The network of claim 4 wherein said pre-calculated delay period exceeds the round-trip delay between said core node and said respective edge node.
6. The network of claim 4 wherein each of said master controllers includes a burst scheduler.
- 5 7. The network of claim 6 wherein said burst scheduler computes a burst-transfer schedule for a plurality of space switches.
8. The network of claim 7 wherein each of said space switches has a plurality of burst-mode input ports, from each of which individual data bursts are directed to output ports of the space switch, and a plurality of channel-mode input ports, the entire data from each of which  
10 is directed to a respective output port of the space switch.
9. The burst-switching network of claim 1 wherein each of said plurality of master controllers comprises:
- an input interface for receiving upstream control bursts from a source node;
  - an input interface to receive control data from a core-node controller;
  - 15 a burst scheduler for generating a schedule for operation of at least one space switch;
  - an output interface for communicating said schedule to a sink node associated with said source node;
  - a transmitter operative to transmit instructions to a slave controller of each of  
20 said at least one space switch, where said instructions are based on said schedule; and
  - a device to acquire timing data from an upstream control burst.
10. In a bufferless space switch having a plurality of burst-mode input ports and a plurality of output ports, a method of determining a schedule for switching data bursts, over a  
25 designated schedule period T, from said plurality of burst-mode input ports to said plurality of output ports, the method including the step of repetitively employing said schedule for switching data bursts during m consecutive periods, m being an integer greater than zero and each of said consecutive periods is equal to said designated period.

11. The method of claim 10 including the further step of setting  $m$  to exceed the ratio of the time required to compute said schedule and said designated schedule period  $T$ .
12. The method of claim 11 including the further step of computing said schedule every  $m$  consecutive periods.
- 5 13. The method of claim 12 including the further step of generating said schedule for a succession of bursts generated over a period  $T$ .
14. The method of claim 13 including the further step of generating said succession of bursts according to bitrate allocations for burst streams to be switched from a burst-mode input port to an output port.
- 10 15. The method of claim 14 including the further step of refreshing said bitrate allocations periodically every  $m \times T$  interval.
16. In a bufferless space switch having a plurality of burst-mode input ports and a plurality of output ports, a method of determining a schedule for switching data bursts, over each of successive time intervals, each time interval having a duration  $T$ , from said plurality of burst-mode input ports to said plurality of output ports, comprising the steps of: setting the computation period for each of said successive time intervals to an integer multiple  $m$  of the interval  $T$ ; and computing  $m$  successive schedules concurrently.
- 15 17. The method of claim 16 including the further step of setting the value of  $m$  to exceed the time required to compute said schedule for each time interval  $T$  divided by the time interval  $T$ .
- 20 18. The method of claim 17 including the further step of operating at least  $m$  scheduling devices concurrently.
19. The method of claim 18 including the further steps of computing said schedule for burst descriptors generated according to bitrate allocations for each pair of burst-mode input port and output port, and refreshing the bitrate allocations at every interval  $T$ .
- 25 20. A method of computing a burst-switching schedule in a bufferless space switch having a plurality of burst-mode input ports, the method comprising steps of :

- 5
- 10
- 15
- 20
- 25
- a. receiving burst descriptors associated with each of the plurality of burst-mode input ports;
  - b. placing said burst descriptors in bursts queues, at least one queue being associated with each one of said plurality of burst-mode input ports;
  - c. cyclically accessing said burst queues, determining corresponding input free time and selecting a maximum of Q candidate burst descriptors;
  - d. determining free time for output port indicated in each candidate burst descriptor;
  - e. determining the absolute value W of the difference between the output-free time corresponding to each of the Q burst descriptors and said input free time; and
  - f. selecting the candidate burst yielding the least value W.
21. The method of claim 20 including the further step of determining said burst descriptors at edge nodes in a closed-loop burst-transfer-control system.
22. The method of claim 20 including the further step of determining said burst descriptors at core nodes in a closed-loop burst-transfer-control system.
23. The method of claim 22 where each of said burst descriptors is associated with a burst stream and said determining is based on a bitrate allocation for said burst stream.
24. A burst scheduler for a space switch, said space switch having a plurality of input ports and a plurality of output ports, said scheduler including:  
a receiver for receiving burst descriptors and placing each of said burst descriptors in one of a plurality burst-descriptor memories, each of said burst descriptors identifying an input port, an output port, and a burst size;  
an input-state memory for storing next available time of each of said input ports;  
a plurality of output-state memories, each storing next-available time of each of said output ports;

a processing circuit including a scheduler kernel for computing a schedule for burst-transfer across said space switch over a predefined period of time T, said processing circuit operable to

select a number Q of candidate burst descriptors for each input port, where  
Q is an integer greater than zero;

compare corresponding entries in said input-state memory and said plurality of output-state memories for each of said Q candidate burst descriptors and determine a corresponding merit index; and

select one of said Q candidate burst descriptors according to said merit index;

and

a permits buffer for storing said schedule.

25. The burst scheduler of claim 24 wherein said merit index is based on an absolute value of the difference between said corresponding entries.

26. The burst scheduler of claim 24 wherein each of said burst-descriptor memories is operative to store burst descriptors belonging to a subset of burst-switching input ports.

27. The burst scheduler of claim 24 wherein the output-state memories have identical content.

28. The burst scheduler of claim 24 wherein the output-state memories are read concurrently at arbitrary memory addresses.

29. A core node having a plurality of space switches, each space switch having burst-mode input ports and channel-mode input ports, wherein each burst-mode input port switches individual data bursts to respective output ports, and each channel-mode input port has a switched channel connection carrying a succession of data units of any format to a single output port.

30. The core node of claim 29 wherein the search for a channel connection starts from the same space switch and follows the same order for each channel connection, and wherein the

number of channel connections in a space switch is limited below a predetermined upper bound.

31. The core node of claim 29 wherein the burst connections are allocated equitably among the space switches.

5 32. In a core node having a plurality of space switches operated in parallel, each of said plurality of space switches having a plurality of input ports, a plurality of output ports, and a master controller with one said master controller designated to function as a core-node controller, said core node switching burst streams from a plurality of upstream links, each having multiple wavelength channels, to a plurality of downstream links, each having  
10 multiple wavelength channels, a method of confining connections from each upstream link to each downstream link to a small number of space switches, the method comprising the steps of:

receiving a bitrate requirement for each connection;

15 sorting received bitrate requirements associated with each upstream link in a descending order according to bitrate value;

implementing a cyclic allocation of said requirements to corresponding paths of the space switches, retaining a remainder when one of said corresponding paths is exhausted, and determining a progress indicator; and

repeating said cyclic allocation if permitted by said progress indicator.